Journal of Water and Soil
https://jsw.um.ac.ir

Research Article
Vol. 36, No. 6, Feb.-Mar., 2023, p. 711-727

Evaluating Reference Evapotranspiration Using Data Mining Methods and Comparing it with the Results of Water Requirement System in Qazvin Province

A. Sedaghat[1]*, N.A. Ebrahimi Pak[2], A. Tafteh[3], S.N. Hosseini[4]

Received: 05-09-2022
Revised: 03-12-2022
Accepted: 06-12-2022
Available Online: 22-02-2023

How to cite this article:
Sedaghat, A., Ebrahimi Pak, N.A., Tafteh, A., & Hosseini, S.N. (2023). Evaluating Reference Evapotranspiration Using Data Mining Methods and Comparing it with the Results of Water Requirement System in Qazvin Province. *Journal of Water and Soil 36*(6): 711-727. (In Persian with English abstract)
DOI: 10.22067/jsw.2022.78652.1200

Introduction

The accuracy of determining reference evapotranspiration (ET_0) is an important factor in estimating agricultural and garden water requirements. The complexity of the evapotranspiration process and its dependence on meteorological data have made it difficult to accurately estimate this variable. Non-linearity, inherent uncertainty and the need for diverse climatic information in ET_0 estimation have been the reasons that have made researchers interested in data mining methods such as artificial neural network (ANNs), random forest (RF) and support vector machine (SVM). Dos et al. (2020) evaluated the performance of machine learning methods to estimate daily ET_0 with limited meteorological data. Their results showed that machine learning methods estimate ET_0 with high accuracy, even in the absence of some variables. The use of artificial intelligence models in estimating ET_0 with high accuracy has become popular in recent years, but the complexity of these models makes it difficult to apply them to regions with different climatic conditions (Feng and Tian, 2021). Therefore, the aim of this study is to show that different data mining methods are suitable for daily ET_0 estimation, which can reach a comprehensive and simple model with high accuracy by using minimal weather data.

Materials and Methods

In this research, the accuracy of data mining methods in estimating ET0 was evaluated in comparison with the plant water requirement system (FAO-Penman-Monteith standard method). For this purpose, data related to meteorological parameters such as sunshine hour, air temperature, wind speed, and relative humidity air were collected from ten synoptic stations and five climatology stations of Qazvin province in a period of 10 years (1389-1399). The ET0 extracted from the plant water requirement system was calculated based on the Penman-Moanteith method of FAO 56 and on a daily time scale, which is the actual value (measured) with the estimated values obtained by data mining methods (ANNs, RF and SVM) were evaluated. In order to validate the obtained results, the data of each station was divided into two sets of training (two-thirds of data) and testing (one-third of data). Finally, the generalizability of the mentioned methods in estimating ET0 was investigated based on NRMSE, R2, RMSE, MBE, EF and d Criteria.

Results and Discussion

1, 2, 3 and 4- Researcher, Associate Professor, Assistant Professor and Researcher of Department of on Farm Water Management, Soil and Water Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran, respectively.
(*- Corresponding Author Email: Azadehsedaghat65@gmail.com)

The results showed that the ET_0 values of the plant water requirement system have a good correlation with the estimated ET_0 values of ANNs, RF, and SVM methods. In this research, the accuracy of the results of ANNs method was relatively higher than the other two methods. The results of statistical investigations and diagrams showed that ANNs, RF and SVM methods, considering all meteorological parameters (mean air temperature, average relative humidity, sunshine hours and wind speed) as input to the model, in Qazvin synoptic station with altitude 1279 meters and the climatology station of Rajaei power plant with a height of 1318 meters, estimated ET_0 with higher accuracy in both training and testing steps.In the ANNs method, the values of NRMSE and R^2 at Qazvin synoptic station in both training and testing steps are equal to 0.11 and 0.97, respectively, and at Rajaei Power Plant climatology station in both training and testing steps are equal to 0.10 and 0.97, respectively. In this research, the accuracy of estimating the value of ET0 in two ANNs and RF methods is close to each other and higher than the SVM method. On the other hand, the fitting speed of the ANNs method is very long compared to the RF method, and considering all aspects, it can be said that the RF method has a more suitable approach for estimating the ET_0 value. The results of this research showed that the value of ET_0 is not only based on air temperature, but may change under the influence of other factors such as air pollution, and is also strongly influenced by regional conditions such as topography and altitude.

Conclusion

The results of this research, in addition to better investigation of ET0, help to know more influential factors in each region and can be used in regions with similar climatic conditions. For example, in the current study area, it was found that the role of average air temperature is greater than other climatic parameters and has a greater impact on ET0. Therefore, it can be said that increasing the average daily air temperature will increase ET0 and subsequently increase the water requirement of plants. As a result, by using these methods and paying attention to these points, it is possible to avoid water stress and possible reduction of the production.

Keywords: Data mining, Reference evapotranspiration, Water requirement system

Journal of Water and Soil
https://jsw.um.ac.ir

مقاله پژوهشی

جلد ۳۶، شماره ۶، بهمن-اسفند ۱۴۰۱، ص. ۷۲۷-۷۱۱

ارزیابی تبخیرتعرق مرجع با استفاده از روش‌های داده‌کاوی و مقایسه آن با نتایج سامانه نیازآب در استان قزوین

آزاده صداقت[1*]- نیازعلی ابراهیمی پاک[2]- آرش تافته[3]- سیده نرگس حسینی[4]

تاریخ دریافت: ۱۴۰۱/۰۶/۱۴
تاریخ پذیرش: ۱۴۰۱/۰۹/۱۵

چکیده

برآورد دقیق تبخیرتعرق مرجع (ET_0) یکی از عوامل مهم برای محاسبهٔ نیاز آبی و آب مصرفی گیاهان زراعی و باغی است. پیچیدگی فرآیند تبخیرتعرق و وابستگی آن به داده‌های هواشناسی برآورد دقیق این متغیر را دشوار کرده است. ویژگی غیرخطی، عدم قطعیت ذاتی و نیاز به اطلاعات اقلیمی متنوع در برآورد ET_0 از دلایلی بوده‌اند که باعث شده پژوهشگران به‌سوی روش‌های داده‌کاوی همچون شبکه عصبی مصنوعی (ANNs)، جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) روی آورند. در این تحقیق، داده‌های هواشناسی در بازه زمانی ده ساله (۱۳۹۹-۱۳۸۹) از ایستگاه‌های هواشناسی استان قزوین جمع‌آوری شد. ابتدا مقادیر ET_0 در سامانه نیاز آب که از روش پنمن-مانتیث محاسبه شد، استخراج گردید. سپس این مقادیر به‌عنوان مقادیر واقعی (اندازه‌گیری شده) با مقادیر تخمینی بدست آمده با روش‌های داده‌کاوی (ANNs، RF و SVM) ارزیابی شد. جهت اعتبارسنجی نتایج بدست آمده، داده‌های هر ایستگاه به دو مجموعه آموزش (دوسوم داده‌ها) و آزمون (یک‌سوم داده‌ها) تقسیم شدند. نتایج بررسی‌های آماری و دیاگرام نشان دادند، در هر سه روش استفاده شده با در نظر گرفتن تمامی پارامترهای هواشناسی (میانگین دمای هوا، میانگین رطوبت نسبی، ساعت آفتابی و سرعت باد) به‌عنوان ورودی مدل، در ایستگاه سینوپتیک قزوین و ایستگاه کلیماتولوژی نیروگاه رجایی، در هر دو مرحله آموزش و آزمون، ET_0 با دقت بالاتری برآورد شد. همچنین در این تحقیق دقت نتایج روش ANNs نسبت به دو روش دیگر به‌طور نسبی بالاتر بوده است. در هر دو مرحله آموزش و آزمون مقادیر NRMSE و R^2 بدست آمده از روش ANNs، در ایستگاه سینوپتیک قزوین برابر و به ترتیب برابر ۰/۱۱ و ۰/۹۷، و در ایستگاه کلیماتولوژی نیروگاه رجایی برابر و به‌ترتیب برابر ۰/۱۰ و ۰/۹۷ می‌باشد. به‌طورکلی نتایج نشان داد که میانگین دمای هوای روزانه مهمترین پارامتر هواشناسی تأثیرگذار در برآورد ET_0 می‌باشد.

واژه‌های کلیدی: تبخیرتعرق مرجع، داده‌کاوی، سامانه نیاز آب

مقدمه

تبخیرتعرق فرآیند از دست رفتن آب از خاک و سطوح محصول به اتمسفر توسط فرآیندهای تبخیر و تعرق است. تبخیرتعرق یکی از مهم‌ترین اجزای چرخه آب است و عامل کلیدی برای کشاورزی، برنامه‌ریزی آبیاری و منابع آب است. پیش‌بینی دقیق و سریع ET_0 به تجزیه‌وتحلیل تغییرات محیطی کمک می‌کند و برای آبیاری گیاه، انتقال آب آبیاری، مدیریت منابع آب در حوضه‌های رودخانه‌ای، ارزیابی محیط‌زیست، در مقیاس‌های مختلف و هیدرولوژیکی و مدل اکوسیستمی ضروری و بسیار مهم است (Feng and Tian, 2021). پیچیدگی فرآیند تبخیرتعرق و وابستگی آن به متغیرهای هواشناسی، عدم دسترسی به تمامی داده‌های هواشناسی و عدم تعمیم‌پذیری یک مدل برای اقلیم‌های مختلف، برآورد دقیق این متغیر را دشوار کرده

۱، ۲، ۳ و ۴- به‌ترتیب محقق، دانشیار، استادیار و محقق بخش مدیریت آب در مزرعه موسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران

(*- نویسنده مسئول: Email: Azadehsedaghat65@gmail.com)
DOI: 10.22067/jsw.2022.78652.1200

است. به‌طورکلی برآورد ET_0 را می‌توان از طریق روش‌های مستقیم و غیرمستقیم انجام داد (Jing *et al*., 2019). ET_0 را می‌توان به‌طور مستقیم بر اساس روش‌های تراز انرژی و انتقال جرم بخار آب اندازه‌گیری کرد، بااین حال، هزینه زیادی دارد. روش‌های غیرمستقیم برای تخمین ET_0 با توسعه معادلات تجربی مانند Penman-Monteith (PM) آغاز شد (Monteith, 1965). بر اساس روش PM، FAO56 یک روش استاندارد برای تخمین ET_0 منتشر کرد. روش PM ارائه‌شده در FAO56 توسط بسیاری از محققین به‌عنوان روش پایه در محاسبه ET_0 استفاده شده است (Saggi and Jain, 2019, Shiri *et al*., 2019). روش PM علاوه بر دقت بالا در تخمین ET_0، دارای محدودیت‌هایی مانند نیاز داشتن به پارامترهای آب و هوایی زیاد، می‌باشد، از طرفی ایستگاه هواشناسی کمی وجود دارد که همه این پارامترها را باهم ارائه دهد (Panaitescu *et al*., 2014). برای غلبه بر این مشکل، بسیاری از محققان مدل‌های تجربی را توسعه داده‌اند که نیازهای کمتری به داده‌های آب و هوایی دارد.

رویکردهای دیگری که اخیراً توجه محققین را به خود جلب کرده است، روش‌های مبنی بر هوش مصنوعی مانند شبکه عصبی مصنوعی (ANNs)[1] (Gavili *et al*., 2018)، ماشین بردار پشتیبان (SVM)[2] (Tabari *et al*., 2013)، و درخت‌های تصمیم‌گیری مانند (RF)[3] (Raza *et al*., 2020) می‌باشد، که کاربردهای زیادی در تخمین ET_0 داشته‌اند. با ظهور فناوری یادگیری ماشین و هوش مصنوعی، محققان شروع به کشف چگونگی ترکیب الگوریتم‌های هوشمند با روش های تخمین سنتی برای تخمین دقیق‌تر و مؤثرتر ET_0 کردند. یونس و همکاران (Üneş *et al*., 2020)، ET_0 روزانه را بر اساس شرایط آب و هوایی با استفاده از معادلات تجربی و روش‌های مختلف داده‌کاوی پیش‌بینی کردند. با توجه به نتایج آن‌ها، معادله مبتنی برتابش بهتر از سایر معادلات تجربی و بالاترین ضریب همبستگی برای روش شبکه عصبی و حداقل خطاها برای تابع مبنای شعاعی SVM به‌دست‌آمده است. چن و همکاران (Chen *et al*., 2020)، ET_0 روزانه را بر اساس داده‌های هواشناسی محدود با استفاده از سه روش یادگیری عمیق، دو روش یادگیری ماشین و هفت معادله تجربی برآورد کردند. نتایج آن‌ها نشان داد که وقتی ویژگی های مبتنی بر دما در دسترس باشد، مدل های یادگیری عمیق به‌طور قابل‌توجهی بهتر از مدل‌های تجربی مبتنی بر دما عمل می‌کنند. دوس سانتوس فاریاس و همکاران (dos Santos Farias *et al*., 2020) عملکرد روش‌های یادگیری ماشین برای تخمین ET_0 روزانه در کشور برزیل ارزیابی کردند. نتایج آن‌ها نشان داد که روش‌های یادگیری ماشین، حتی در غیاب برخی متغیرها، ET_0 را با دقت بالا برآورد می‌کنند. متار (Mattar, 2018) هشت مدل برای تخمین ET_0 با استفاده از الگوریتم ژنتیک ایجاد کرد. نتایج این تحقیق نشان داد، نتایج برآورد بسیار نزدیک به تخمین با مدل FAO-PM-56 بوده است.

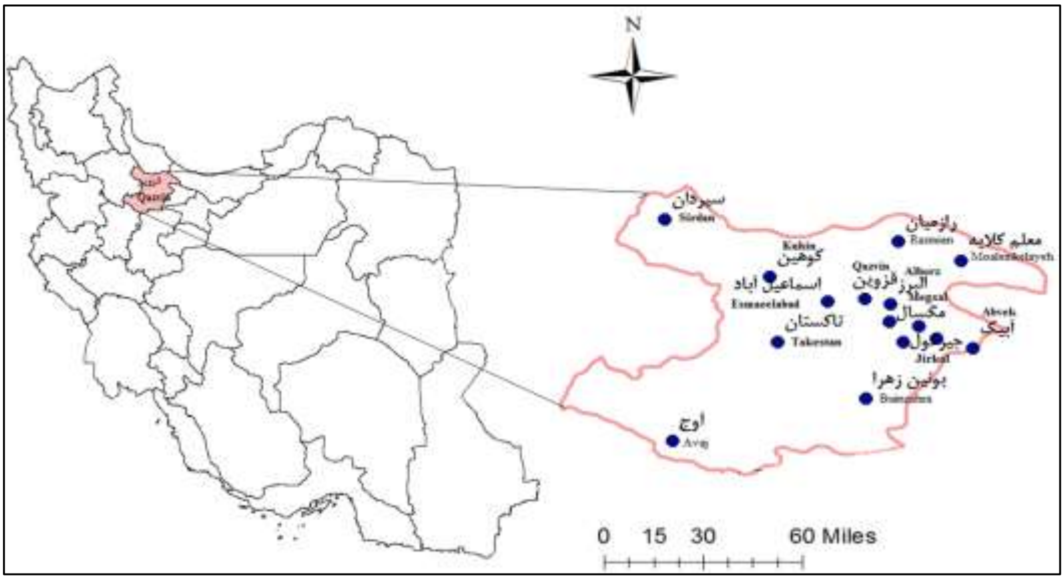
استفاده از روش های هوش مصنوعی در تخمین ET_0 با دقت بالا در سال‌های اخیر رواج یافته است، اما پیچیدگی این روش‌ها و نیاز آن‌ها به داده‌های هواشناسی بیشتر، کاربرد آن‌ها را برای مناطق مختلف دشوار می‌کند (Feng and Tian, 2021). بنابراین، در مقالهٔ حاضر ET_0 از چهار متغیر هواشناسی (میانگین دمای هوای (T_{mean})[4]، میانگین رطوبت نسبی (RH_{mean})[5]، سرعت وزش باد در ارتفاع دو متری (W.S)[6] و ساعت آفتابی (*Sun*)[7]) که از ده ایستگاه هواشناسی سینوپتیک و پنج ایستگاه کلیماتولوژی استان قزوین به‌دست‌آمده است، با استفاده از روش‌های داده‌کاوی تخمین زده شد. هدف از این مطالعه نشان دادن شایستگی روش‌های داده‌کاوی جهت رسیدن به یک مدل جامع و ساده، و معتبر با حداقل داده‌های آب و هوایی برای تخمین ET_0 روزانه می‌باشد.

مواد و روش‌ها

موقعیت محدوده موردمطالعه

استان قزوین در حوزه مرکزی ایران بین ۴۸ درجه و ۴۵ دقیقه تا ۵۰ درجه و ۵۱ دقیقه طول شرقی از نصف‌النهار گرینویچ و ۳۵ درجه و ۲۴ دقیقه تا ۳۶ درجه و ۴۸ دقیقه عرض شمالی نسبت به خط استوا قرار دارد (شکل ۱). استان قزوین یکی از استان‌های مرتفع ایران است و ارتفاع بیش از ۹۰ درصد سطح استان بیش از متوسط ارتفاع (۱۲۰۰متر) است. اختلاف شدید ارتفاع و وجود ناهمواری از عوامل اصلی تنوع آب‌وهوای استان قزوین است. بیش از نیمی از استان قزوین یعنی ۵۳ درصد دارای اقلیم خشک فراسرد تا سرد و ۱۸ درصد از سطح استان دارای اقلیم مدیترانه‌ای فراسرد تا سرد می‌باشد. در این پژوهش از داده‌های مربوط به دوره ده سالهٔ (۱۳۸۹-۱۳۹۹) استفاده شد.

1- Artificial Neural Networks
2- Support Vector Mashine
3- Random Forest
4- Mean temperature
5- Mean relative humidity
6- Wind Speed
7- Sunshine



شکل ۱- موقعیت جغرافیایی استان قزوین

Figure 1- Geographical location of Qazvin province

برآورد مقدار ET_0 سامانه نیاز آب گیاهان

مقدار ET_0 در این پژوهش از بانک اطلاعاتی سامانه نیاز آب مؤسسه تحقیقات خاک و آب کشور که به‌منظور تعیین نیاز آبی استاندارد و واقعی گیاهان باغی و زراعی طراحی‌شده است، استفاده شد (http://niwr.ir). استفاده از این سامانه به‌این‌علت می‌باشد که در حال حاضر دانش فنی جامعی که بتواند از بانک‌های اطلاعات کلان کشور برای برآورد و تعیین نیاز آبی، آب مصرفی و برنامه‌ریزی آبیاری در مناطق مختلف، استفاده شود، وجود ندارد. این سامانه بانک اطلاعاتی خوبی می‌باشد که در زمان کم، حجم زیادی از اطلاعات منطقه را در اختیار کاربران قرار می‌دهد. مقدار ET_0 استخراج‌شده از سامانه نیاز آب بر اساس روش پنمن-مانتیث فائو ۵۶[1] و در مقیاس زمانی روزانه محاسبه‌شده است، که به‌عنوان مقدار واقعی (اندازه‌گیری شده) در مدل‌سازی با روش‌های داده‌کاوی در نظر گرفته شد.

توصیفی از مراحل تخمین ET_0 و روش‌های مدلسازی

برای ایجاد مدل در تمام ایستگاه‌های هواشناسی، هم‌راستایی بین متغیرهای مستقل (فرض صحت رگرسیون خطی) از طریق آماره عامل تورم واریانس (VIF)[2] و با استفاده از نرم‌افزار SPSS.26 بررسی شد. متغیرهایی که VIF آن‌ها کمتر از ۱۰ باشد مشکل هم‌راستایی نخواهند داشت (Hocking, 2013). جهت اعتبارسنجی نتایج به‌دست‌آمده از این تحقیق، داده‌های هر ایستگاه به دو مجموعه آموزش (دوسوم داده‌ها) و آزمون (یک‌سوم داده‌ها) تقسیم شدند.

در این پژوهش از روش‌های شبکه عصبی مصنوعی (ANNs)، ماشین بردار پشتیبان (SVM) و جنگل تصادفی (RF) جهت مدل‌سازی ET_0 استفاده شد. در این پژوهش برای مدل‌سازی ET_0 با روش‌های ANNs، SVM و RF از نرم‌افزار STATISTICA12 استفاده شد. شبکه عصبی پرسپترون چند لایه (MLP)[3] و تابع پایه‌ی شعاعی (RBF)[4]، دو نوع شبکه از روش ANNs می‌باشد که در این پژوهش مورد استفاده قرار گرفت (Hagan *et al.*, 1996; Picton, 2000). روش ANNs در ۱۰۰ تکرار جهت کاهش خطا در مدل‌سازی، تنظیم شد. میانگین ۵۰ خروجی از ۱۰۰ تکرار که خطای کمتری دارند، به‌عنوان خروجی نهایی برای تخمین ET_0 انتخاب شد. روش SVM دارای عامل‌های تنظیم‌کننده شامل C و nu هسته می‌باشد که دارای چهار نوع کرنل شامل کرنل خطی، چندجمله‌ای، شعاعی و حلقوی است (Vapnik, 2000). روش RF یک روش جدید و قدرتمند است که پیشرفت‌های فراوانی را در فنآوری داده‌کاوی داشت. اساس این روش توسط (Breiman, (2001 ارائه شده است.

معیارهای ارزیابی دقت و قابلیت اعتماد مدلها

برای بررسی دقت و قابلیت اعتماد مدل‌های ایجادشده از

1- FAO-56 Penman-Monteith
2- Variance Inflation Factor
3- Multilayer perceptron's
4- Radial basis function

معیارهای آماره‌های ضریب تعیین (R^2)[۱]، مجذور میانگین مربعات خطا (RMSE[۲])، و میانگین انحراف خطای (MBE[۳])، مجذور میانگین خطای نرمال شده (NRMSE) و ضریب کارایی مدل (EF) بهره‌گیری شد.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \quad (۱)$$

$$R^2 = 1 - \frac{\sum_{1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{1}^{N}\left(\hat{y}_i - \bar{y}_i\right)^2} \quad (۲)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i) \quad (۳)$$

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2}}{\bar{X}} \quad (۴)$$

$$d = 1 - \left[\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(|x_i - \bar{X}| + |y_i - \bar{Y}|)^2}\right] \quad (۵)$$

$$EF = 1 - \frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(x_i - \bar{X})^2} \quad (۶)$$

از لحاظ آماری حداقل مقدار RMSE زمانی مقادیر تخمینی با مقادیر اندازه‌گیری برابر صفر باشد، مدل دارای دقت بیشتری خواهد بود (Adab *et al*., 2020). MBE عمدتاً برای نشان دادن اریب در مدل، استفاده می‌شود. وقتی مقدار آن مثبت باشد نشان‌دهندهٔ بیش برآوردی و مقدار منفی آن نشان‌دهندهٔ کم‌برآوردی تخمین می‌باشد (Gopinathan, 1988). از NRMSE به‌عنوان شاخص خطای پیش‌بینی مدل استفاده می‌شود، اگر مقدار آن کمتر از ۱۰٪ باشد، نتیجه تخمین عالی، بین ۱۰ تا ۲۰٪ ، خوب؛ بین ۲۰ تا ۳۰٪ ، قابل قبول؛ و برای مقادیر بیشتر از ۳۰٪، ضعیف می‌باشد (Sandhu and Irmak, 2020). شاخص d، شاخص توافق ویلموت می‌باشد که از دامنه صفر (بدون توافق) تا یک (توافق کامل) می‌باشد که میزان درستی و صحت برآوردها را نشان می‌دهد (Willmott *et al*., 2012). همچنین مقدار EF از منفی بی‌نهایت تا یک به ترتیب برای بدترین و بهترین برآورد تغییرمی‌کند و کفایت و کارآیی مدل را نشان می‌دهد و بالای ۰/۵ عملکرد قابل‌قبول مدل را نشان می‌دهد (Karimipour and Banitalebi, 2020).

1- Coefficient of determination
2- Root mean square error
3- Mean Bias Error

نتایج و بحث

توصیف آماری

ویژگی‌های آماری متغیرهای موردمطالعه در همه ایستگاه‌های سینوپتیک و کلیماتولوژی در جدول ۱ آورده شده است. تمامی داده‌های ورودی و خروجی در همه ایستگاه‌ها در این تحقیق، دارای دامنه تغییرات بالایی (اختلاف بین کمترین و بیشترین مقدار) می‌باشند، که نشان‌دهندهٔ این است که باعث افزایش قابلیت تعمیم‌پذیری مدل می‌شود (جدول ۱). به‌منظور نشان دادن پراکندگی نسبی داده‌ها به میانگین، از ضریب تغییرات (CV)استفاده شد. همبستگی بین عامل‌های ورودی مدل با ET_0 بررسی شد (جدول ۲). نتایج همبستگی نشان داد، ET_0 در همه ایستگاه‌ها با ساعت آفتابی، سرعت باد، میانگین دمای هوا همبستگی مثبت و معنی‌دار، و با میانگین رطوبت نسبی همبستگی منفی و معنی‌داری دارد (جدول ۲). با افزایش تابش خورشید و بالا رفتن دمای هوا و همچنین افزایش سرعت باد، مقدار ET_0 افزایش‌یافته و با افزایش رطوبت هوا مقدار ET_0 کاهش می‌یابد (Ndiaye *et al*., 2020).

ویژگی‌های روش مورداستفاده در این پژوهش

شبکه عصبی پرسپترون چند لایه (MLP) و تابع پایه‌ی شعاعی (RBF)، دو نوع شبکه از روش ANNs می‌باشد که در این پژوهش در ایجاد مدل مورداستفاده قرار گرفت (Picton, 2000). در نتایج بدست آمده از روش ANNs، در همهٔایستگاه‌ها، تعداد شبکه MLP استفاده‌شده بیشتر از شبکه RBF می‌باشد. بنابراین شبکه MLP نسبت به RBF در برآورد مقدار ET_0 از عملکرد بهتری برخوردار است. در روش ANNs از توابع فعال‌سازی مختلفی در لایه‌های میانی و خروجی استفاده شد، که Exponential و Tanh تابع فعال‌سازی غالب در لایه پنهان و Identity تابع فعال‌سازی غالب در لایه خروجی بوده و بهترین تابع برای توصیف روابط بین متغیرهای ورودی و ET_0 می‌باشند (جدول ۳). یکی از مهم‌ترین ویژگی روش RF، ساده‌سازی مدل از طریق تعیین اهمیت عامل[۴](تخمین‌گرها) می‌باشد (جدول ۴).

4- Important parameter

جدول ۱- عوامل آماری متغیرهای مورد مطالعه

Table 1- Statistical factors of the studied variables

نام ایستگاه Station name	عامل‌های اقلیمی Climatic factors	حداقل Min	حداکثر Max	میانگین Mean	انحراف استاندارد SD*	ضریب تغییرات CV	نام ایستگاه Station name	عامل‌های اقلیمی Climatic factors	حداقل Min	حداکثر Max	میانگین Mean	انحراف استاندارد SD*	ضریب تغییرات CV
ایستگاه‌های سینوپتیک Synoptic stations													
رازمیان Razmian (n=2192)	میانگین دمای هوا T_{mean} (°C)	-6.8	35.9	17.1	8.9	0.5	سیردان Sirdan (n=2192)	میانگین دمای هوا T_{mean} (°C)	-10.1	37.0	13.2	8.5	0.6
	میانگین رطوبت نسبی RH_{mean} (%)	0.00	98	55.80	16.89	0.30		میانگین رطوبت نسبی RH_{mean} (%)	0.00	100	63.43	22.60	0.36
	سرعت باد W.S (ms^{-1})	0.00	5.23	1.35	0.64	0.48		سرعت باد W.S (ms^{-1})	0.00	5.87	1.17	0.76	0.65
	ساعت آفتابی Sun (h)	0.00	12.90	6.50	4.11	0.63		ساعت آفتابی Sun (h)	0.00	14.51	9.11	3.57	0.39
	تبخیرتعرق مرجع ET_0 (mm day^{-1})	0.46	9.45	3.37	2.01	0.60		تبخیرتعرق مرجع ET_0 (mm day^{-1})	0.29	8.54	3.00	1.88	0.63
بوئین زهرا Buinzahra (n=4018)	میانگین دمای هوا T_{mean} (°C)	-7.5	36.0	15.9	9.5	0.6	کوهین Kuhin (n=4018)	میانگین دمای هوا T_{mean} (°C)	-12.8	40.5	12.7	9.1	0.7
	میانگین رطوبت نسبی RH_{mean} (%)	6.50	100	49.02	18.29	0.37		میانگین رطوبت نسبی RH_{mean} (%)	10.50	100	61.26	20.06	0.33
	سرعت باد W.S (ms^{-1})	0.00	4.30	1.60	0.63	0.40		سرعت باد W.S (ms^{-1})	0.74	7.92	1.89	0.85	0.45
	ساعت آفتابی Sun (h)	0.00	13.85	8.62	3.44	0.40		ساعت آفتابی Sun (h)	0.00	13.20	7.49	3.69	0.49
	تبخیرتعرق مرجع ET_0 (mm day^{-1})	0.41	9.78	3.88	2.23	0.57		تبخیرتعرق مرجع ET_0 (mm day^{-1})	0.29	12.15	3.27	2.01	0.62
تاکستان Takestan (n=4018)	میانگین دمای هوا T_{mean} (°C)	-9.10	34.80	14.87	9.23	0.6	اسماعیل آباد Esmaeel Abad (n=4018)	میانگین دمای هوا T_{mean} (°C)	-9.7	34.9	15.0	9.4	0.6
	میانگین رطوبت نسبی RH_{mean} (%)	14.50	100.00	51.58	17.34	0.34		میانگین رطوبت نسبی RH_{mean} (%)	0.00	99.50	53.66	17.79	0.33
	سرعت باد W.S (ms^{-1})	0.00	9.72	1.72	0.85	0.49		سرعت باد W.S (ms^{-1})	0.00	3.73	1.39	0.45	0.32
	ساعت آفتابی Sun (h)	0.00	13.50	7.69	3.67	0.48		ساعت آفتابی Sun (h)	0.00	13.10	8.20	3.62	0.44
	تبخیرتعرق مرجع ET_0 (mm day^{-1})	0.37	10.56	3.65	2.19	0.60		تبخیرتعرق مرجع ET_0 (mm day^{-1})	0.41	8.95	3.61	2.20	0.61
قزوین Qazvin (n=4018)	میانگین دمای هوا T_{mean} (°C)	-8.15	34.00	15.11	9.22	0.6	آبیک Abyek (n=4018)	میانگین دمای هوا T_{mean} (°C)	-10.8	38.4	15.6	9.5	0.6
	میانگین رطوبت نسبی RH_{mean} (%)	15.00	99.50	53.61	16.60	0.31		میانگین رطوبت نسبی RH_{mean} (%)	0.00	100	53.33	19.63	0.37
	سرعت باد W.S (ms^{-1})	0.40	4.71	1.42	0.42	0.30		سرعت باد W.S (ms^{-1})	0.00	5.16	1.48	0.51	0.35
	ساعت آفتابی Sun (h)	0.00	13.10	8.20	3.62	0.44		ساعت آفتابی Sun (h)	0.00	14.45	8.93	3.43	0.38
	تبخیرتعرق مرجع ET_0 (mm day^{-1})	0.38	8.94	3.64	2.21	0.61		تبخیرتعرق مرجع ET_0 (mm day^{-1})	0.39	9.63	3.82	2.30	0.60

معلم کلایه Moalemkelayeh (n=4018)	میانگین دمای هوا T_{mean} (°C)	-9.95	33.3	13.8	8.8	0.6	اوج Avaj (n=4018)	میانگین دمای هوا T_{mean} (°C)	-14.0	30.0	10.9	9.1	0.8
	میانگین رطوبت نسبی RH_{mean} (%)	13.50	100	57.70	18.68	0.32		میانگین رطوبت نسبی RH_{mean} (%)	1.00	100	51.11	19.74	0.39
	سرعت باد W.S (ms^{-1})	0.00	3.92	1.49	0.55	0.37		سرعت باد W.S (ms^{-1})	0.00	8.37	1.54	0.62	0.40
	ساعت آفتابی Sun (h)	0.00	13.00	6.42	3.92	0.61		ساعت آفتابی Sun (h)	0.00	13.40	5.23	4.14	0.79
	تبخیرتعرق مرجع ET_o (mm day^{-1})	0.34	8.23	3.17	1.92	0.60		تبخیرتعرق مرجع ET_o (mm day^{-1})	0.35	8.98	2.95	1.73	0.59
نام ایستگاه	عامل‌های اقلیمی	Min	Max	Mean	SD	CV	نام ایستگاه	عامل‌های اقلیمی	Min	Max	Mean	SD	CV
ایستگاه کلیماتولوژی Climatology stations													
باغ کوثر Bagh kosar (n=4018)	میانگین دمای هوا T_{mean} (°C)	-9.25	35.70	15.46	9.17	0.59	جیرکول Jirkol (n=4018)	میانگین دمای هوا T_{mean} (°C)	-9.25	36.55	15.52	9.12	0.59
	میانگین رطوبت نسبی RH_{mean} (%)	10.50	100	60.98	15.62	0.26		میانگین رطوبت نسبی RH_{mean} (%)	7.50	100	59.51	18.49	0.31
	سرعت باد W.S (ms^{-1})	0.00	2.50	0.81	0.63	0.78		سرعت باد W.S (ms^{-1})	0.00	3.55	0.83	0.62	0.74
	ساعت آفتابی Sun (h)	0.53	14.26	9.47	3.02	0.32		ساعت آفتابی Sun (h)	0.00	14.26	8.99	3.05	0.34
	تبخیرتعرق مرجع ET_o (mm day^{-1})	0.42	9.31	3.37	2.13	0.63		تبخیرتعرق مرجع ET_o (mm day^{-1})	0.42	9.30	3.29	2.05	0.62
مگسال Megsal (n=4018)	میانگین دمای هوا T_{mean} (°C)	-9.25	35.70	15.46	9.17	0.59	البرز Alborz (n=4018)	میانگین دمای هوا T_{mean} (°C)	-9.70	34.85	15.14	9.41	0.62
	میانگین رطوبت نسبی RH_{mean} (%)	10.50	100	60.98	15.62	0.26		میانگین رطوبت نسبی RH_{mean} (%)	11.50	99.50	53.27	17.62	0.33
	سرعت باد W.S (ms^{-1})	0.00	2.50	0.81	0.63	0.78		سرعت باد W.S (ms^{-1})	0.40	2.50	1.42	0.41	0.29
	ساعت آفتابی Sun (h)	0.53	14.26	9.47	3.02	0.32		ساعت آفتابی Sun (h)	0.00	13.10	8.23	3.58	0.44
	تبخیرتعرق مرجع ET_o (mm day^{-1})	0.42	9.31	3.37	2.14	0.63		تبخیرتعرق مرجع ET_o (mm day^{-1})	0.43	8.95	3.65	2.22	0.61
نیروگاه رجایی Rajai power plant (n=4018)	میانگین دمای هوا T_{mean} (°C)	-9.25	35.70	15.47	9.10	0.59							
	میانگین رطوبت نسبی RH_{mean} (%)	0.00	100	51.38	16.89	0.33							
	سرعت باد W.S (ms^{-1})	0.00	6.82	1.05	0.46	0.44							
	ساعت آفتابی Sun (h)	0.00	14.09	9.10	3.08	0.34							
	تبخیرتعرق مرجع ET_o (mm day^{-1})	0.37	8.74	3.47	2.07	0.60							

SD: انحراف استاندارد

SD: Standard Deviation

جدول ۲- رابطه همبستگی بین عامل‌های هواشناسی با مقدار ET_0

Table 2- Correlation between meteorological factors and ET_0

	ایستگاه‌ها Stations	میانگین دمای هوا T_{mean} (°C)	میانگین رطوبت نسبی RH_{mean} (%)	سرعت باد W.S (ms^{-1})	ساعت آفتابی Sun (h)	ایستگاه‌ها Stations	میانگین دمای هوا T_{mean} (°C)	میانگین رطوبت نسبی RH_{mean} (%)	سرعت باد W.S (ms^{-1})	ساعت آفتابی Sun (h)
ETo	آبیک Abyek	0.92**	-0.69**	0.46**	0.70**	کوهین Kuhin	0.91**	-0.60**	0.29**	0.53**
ETo	اسماعیل آباد Esmaeel Abad	0.93**	-0.60**	0.46**	0.71**	معلم کلایه Moalemkelayeh	0.91**	-0.53**	0.45**	0.57**
ETo	آوج Avaj	0.89**	-0.72**	0.11**	0.47**	البرز Alborz	0.93**	-0.72**	0.48**	0.70**
ETo	بویین زهرا Buinzahra	0.93**	-0.78**	0.21**	0.56**	باغ کوثر Bagh kosar	0.92**	-0.75**	0.39**	0.81**
ETo	تاکستان Takestan	0.91**	-0.75**	0.39**	0.50**	جیرکول Jirkol	-0.64**	0.89**	0.42**	0.73**
ETo	رازمیان Razmian	0.88**	-0.51**	0.52**	0.48**	مگسال Megsal	0.92**	-0.75**	0.39**	0.81**
ETo	سیردان Sirdan	0.87**	-0.35**	0.44**	0.54**	نیروگاه رجایی Rajai power plant	0.90**	-0.75**	0.30**	0.84**
ETo	قزوین Qazvin	0.93**	-0.72**	0.45**	0.71**	کوهین Kuhin	0.91**	-0.60**	0.29**	0.53**

*و** به ترتیب همبستگی معنی‌دار در سطوح ۵ و ۱ درصد را نشان می‌دهند.

* and** Significant correlation at levels of 5 and 1 percent, respectively

در جدول ۴ مشاهده می‌شود در همه ایستگاه‌ها به‌جز جیرکول میانگین دمای هوا بیشترین سهم و مهم‌ترین مؤلفه را در افزایش دقت تخمین ET_0 یا به‌عبارت‌دیگر بیشترین سهم را در کاهش خطای تخمین ET_0 دارد. ماجوزی و همکاران (Majozi *et al.*, 2021) نتایج مشابهی را ارائه نمودند. هم‌چنین سرعت باد کمترین سهم را در افزایش دقت تخمین مقدار ET_0 دارد. در این تحقیق از رگرسیون بردار پشتیبان نوع دوم استفاده‌شده است. رگرسیون بردار پشتیبان دارای پارامترهای تنظیم C و nu و پارامترهای هسته (γ) است. در این پژوهش نوع کرنل، از هسته تابع پایه شعاعی (RBF) استفاده‌شده است که این هسته شامل پارامتر گاما (γ) است.

با توجه به جدول ۵، نتایج مقایسه سه روش در برآورد مقدار ET_0، نشان داد که در هر دو مرحله آموزش و آزمون دقت روش ANNs نسبت به دو روش RF و SVM به‌طور نسبی بهتر بود.

مقایسه میانگین نتایج حاصل از سه روش بر اساس آماره NRMSE در ایستگاه‌های سینوپتیک و کلیماتولوژی استان قزوین در مرحله آزمون، در شکل ۲ ارائه‌شده است. در شکل ۲ مشاهده می‌شود در ایستگاه‌های سینوپتیک نتایج روش RF و SVM مشابه و با نتایج روش ANNs تفاوت معنی‌داری دارند. در ایستگاه‌های کلیماتولوژی نتایج سه روش تفاوت معنی‌داری ندارند ولی میانگین خطای دو روش RF و ANNs تقریباً مشابه و کمتر از روش SVM می‌باشند.

نتایج نشان می‌دهد (جدول ۵) که میانگین خطای سه روش به‌کاررفته در این پژوهش در بوئین‌زهرا، قزوین و تاکستان که تقریباً ارتفاع یکسانی نسبت به سایر ایستگاه‌های سینوپتیکی دارند، در هر دو مرحله آموزش و آزمون کمتر می‌باشد. افزایش دما، افزایش ET_0 و تغییرات زمانی بارش در دشت‌ها باعث کاهش منابع آب و افزایش نیاز آبی گیاهان شده و در صورت تصمیم‌گیری نامناسب برای سازگاری با شرایط آتی باعث خسارات جبران‌ناپذیر خواهد شد (Fazeli Khiavi *et al.*, 2020). بنابراین با توجه به اینکه این سه منطقه به‌صورت دشت هستند و میانگین دما، سرعت باد و ET_0 در این ایستگاه‌ها نسبت به سایر ایستگاه‌ها بالاتر می‌باشد (جدول ۱)، بنابراین جهت تأمین نیاز آبی گیاهان این مناطق نیاز به تصمیم‌گیری مناسب جهت مدیریت منابع آب می‌باشد. در جدول ۱ مشاهده می‌شود، ضریب تغییرات دما برای ایستگاه‌های بوئین زهرا، قزوین و تاکستان برابر ۰/۶ و ضریب تغییرات ET_0 برای ایستگاه‌های بوئین‌زهرا، قزوین و تاکستان به ترتیب ۰/۵۷، ۰/۶۰ و ۰/۶۱ می‌باشد. با توجه به یکسان بودن ضریب تغییرات دما و ET_0 و همچنین در ارتفاع تقریباً یکسانی از سطح دریا قرار دارند، دقت تخمین ET_0 در این سه ایستگاه تقریباً یکسان می‌باشد (شکل ۲).

جدول ۳- ویژگی‌های روش ANNs بکار رفته در این پژوهش

Table 3- Characteristics of ANNs method used in this study

Stations	نوع شبکه* Network type		تعداد سلول‌های پنهان MLP Number of hidden cells of MLP	تعداد سلول‌های پنهان RBF Number of hidden cells RBF of	توابع لایه‌های خروجی Output layer function					توابع لایه‌های پنهان Hidden layer function					
	MLP	RBF	Range	Range	Tanh	Identity	Sine	Exponential	Logistic	Gaussian	Identity	Sine	Exponential	Tanh	Logistic
آبیک Abyek	47	3	3-10	24-29	9	14	11	8	8	3	7	11	5	16	8
اسماعیل آباد Esmaeel Abad	41	9	3-10	21-29	7	24	5	7	7	9	8	4	9	9	11
البرز Alborz	38	12	3-10	21-30	7	21	8	9	5	12	3	6	11	5	13
آوج Avaj	48	2	3-10	21-25	9	13	7	9	12	2	6	4	14	13	11
باغ کوثر Bagh kosar	49	1	3-10	29	9	18	6	6	11	1	11	7	10	13	8
بویین زهرا Buinzahra	46	4	3-10	21-30	9	16	4	10	11	4	10	4	9	9	14
تاکستان Takestan	48	2	3-10	23-29	7	13	4	10	16	2	5	3	8	15	17
جیرکول Jirkol	42	8	3-9	14-29	9	15	9	9	8	8	6	6	10	11	9
رازمیان Razmian	45	5	3-10	22-28	14	17	6	4	9	5	3	8	13	11	10
سیردان Sirdan	31	19	3-10	21-30	5	27	6	6	6	19	0	4	6	6	15
قزوین Qazvin	47	3	3-10	22-27	6	18	8	10	8	3	7	5	13	12	10
کوهین Kuhin	47	3	3-10	24-30	11	9	10	11	9	3	2	3	15	18	9
معلم کلایه Moalemkelayeh	43	7	2-8	15-25	10	13	9	10	8	7	4	6	9	13	11
مکسال Megsal	48	2	3-10	22-23	8	9	9	11	13	2	5	13	12	11	7
نیروگاه رجایی Rajai power plant	50	0	3-10	0	9	11	9	9	12	0	9	11	10	8	12

*در این ستون نوع و تعداد شبکه استفاده شده از ۵۰خروجی شبکه را نشان می‌دهد.

* The network type and number of outputs used are shown in this column.

جدول ۴- اهمیت عامل به‌دست‌آمده با روش RF

Table 4- Importance of the factor obtained by RF method

ایستگاه‌ها Stations	میانگین دما T_{mean}	میانگین رطوبت نسبی RH_{mean}	سرعت باد W.S	ساعت آفتابی Sun	ایستگاه‌ها Stations	میانگین دما T_{mean}	میانگین رطوبت نسبی RH_{mean}	سرعت باد W.S	ساعت آفتابی Sun
ایستگاه‌های سینوپتیک Synoptic stations					ایستگاه‌های کلیماتولوژی Climatology stations				
آبیک Abyek	1.00	0.61	0.37	0.65	البرز Alborz	1.00	0.53	0.32	0.68
اسماعیل آباد Esmaeel Abad	1.00	0.40	0.34	0.70	باغ کوثر Bagh kosar	1.00	0.78	0.34	0.95
آوج Avaj	1.00	0.58	0.19	0.57	جیرکول Jirkol	0.57	1.00	0.36	0.80
بوئین زهرا Buinzahra	1.00	0.62	0.18	0.53	مگسال Megsal	1.00	0.80	0.35	0.98
تاکستان Takestan	1.00	0.31	0.57	0.43	نیروگاه رجایی Rajai power plant	1.00	0.80	0.43	0.99
رازمیان Razmian	1.00	0.47	0.45	0.56					
سیردان Sirdan	1.00	0.24	0.39	0.58					
قزوین Qazvin	1.00	0.59	0.34	0.74					
کوهین Kuhin	1.00	0.42	0.24	0.46					
معلم کلایه Moalemkelayeh	1.00	0.39	0.37	0.49					

با توجه به جدول ۵ و شکل ۳، دو ایستگاه رازمیان و سیردان با وجود اینکه ارتفاع از سطح دریای آن‌ها نسبت به سایر ایستگاه‌های سینوپتیک کمتر می‌باشد، اما در هر دو مرحله آموزش و آزمون NRMSE آن‌ها بیشتر از سایر ایستگاه‌ها است، و درنتیجه دقت تخمین ET_0 در این ایستگاه پایین‌تر است. ازیک‌طرف تعداد داده‌های هواشناسی در دسترس در این دو ایستگاه (n=2192) محدود بوده است. از طرف دیگر با توجه به نتایج حاصل از تحقیقات شهریاری و همکاران (Shahryar *et al*., 2019) که ایستگاه‌های سینوپتیک استان قزوین را از نظر عوامل اقلیمی بررسی کردند، مشاهده شد بیشترین درصد رطوبت نسبی (۶۰ درصد) را ایستگاه سیردان و کمترین درصد رطوبت نسبی را بوئین‌زهرا (۴۴ درصد) و تاکستان (۴۸ درصد) دارند و همچنین بیشترین سرعت متوسط باد سالانه را ایستگاه‌های سیردان و کوهین دارند.

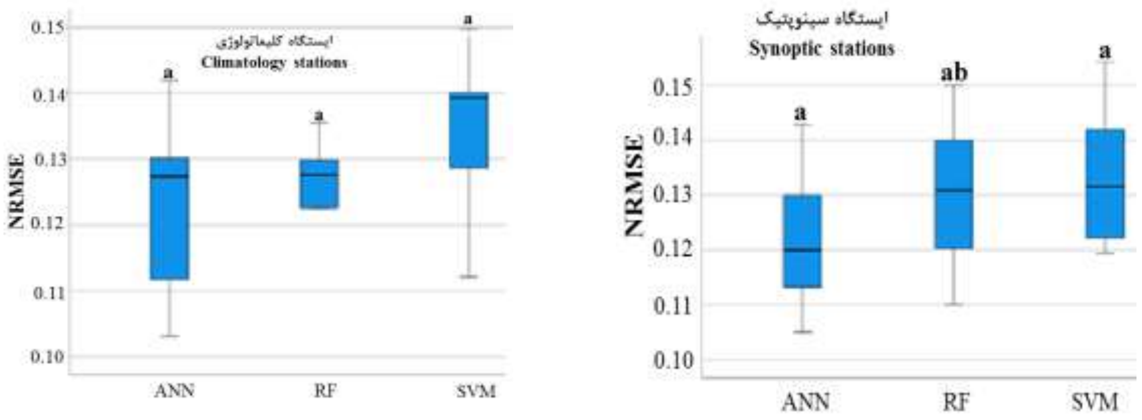
با توجه به جدول ۱، ضریب تغییرات سرعت باد در ایستگاه سیردان (۰/۶۵) بالا می‌باشد. بنابراین احتمالاً به علت شرایط اقلیمی ویژه‌ای که در سیردان و رازمیان وجود دارد و همچنین محدود بودن داده‌های در دسترس، مقدار تخمین ET_0 در این ایستگاه‌ها پایین می‌باشد (شکل ۳). مقادیر آماره‌های RMSE، R^2 و NRMSE روش ANNs در سیردان، به ترتیب برای هر دو مرحله آموزش و آزمون برابر با ۰/۴۳، ۰/۹۵، و ۰/۱۴ می‌باشد. درروش RF مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۴۵، ۰/۹۴ و ۰/۱۵، و برای آزمون برابر با ۰/۴۷، ۰/۹۳، و ۰/۱۶ می‌باشد. درروش SVM مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۴۶، ۰/۹۴ و ۰/۱۵، و برای آزمون برابر با ۰/۴۷، ۰/۹۴، و ۰/۱۵ می‌باشد (جدول ۵).

با توجه به جدول ۵، اگر همه آماره‌های بکار رفته در این تحقیق را در نظر بگیریم، در هر دو مرحله آموزش و آزمون بیشترین دقت تخمین ET_0 در ایستگاه قزوین مشاهده می‌شود. در ایستگاه قزوین، درروش ANNs مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۴۰، ۰/۹۷ و ۰/۱۱، و برای آزمون برابر با ۰/۴۰، ۰/۹۷، و ۰/۱۱ می‌باشد. درروش RF مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۴۲، ۰/۹۶ و ۰/۱۲، و برای آزمون برابر با ۰/۴۳، ۰/۹۶، و ۰/۱۲ می‌باشد. درروش SVM مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۴۴، ۰/۹۶ و ۰/۱۲، و برای آزمون برابر با ۰/۴۴، ۰/۹۶، و ۰/۱۲ می‌باشد (جدول ۵). با توجه به شکل ۳ و ۴، مقادیر NRMSE در ایستگاه‌های کلیماتولوژی کمتر از سینوپتیک می‌باشد، بنابراین دقت تخمین ET_0 ایستگاه‌های کلیماتولوژی بالاتر است.

جدول ۵- نتایج روش‌های مختلف در برآورد مقدار ET_0 در ایستگاه های هواشناسی

Table 5- The results of different methods used for estimating of ET_0 in the meteorological stations

		آموزش (Train)						آزمون (Test)					
ایستگاه‌های سینوپتیک Synoptic stations													
		RMSE	R^2	MBE	NRMSE	d	EF	RMSE	R^2	MBE	NRMSE	d	EF
آوج Avaj	RF	0.36	0.96	-0.002	0.12	0.99	0.96	0.39	0.95	0.017	0.13	0.99	0.94
	ANNs	0.36	0.96	-0.003	0.12	0.99	0.96	0.36	0.95	0.012	0.12	0.99	0.95
	SVM	0.40	0.95	-0.064	0.13	0.99	0.95	0.39	0.95	-0.058	0.13	0.99	0.94
آبیک Abyek	RF	0.48	0.96	-0.025	0.13	0.99	0.96	0.52	0.95	-0.069	0.13	0.99	0.95
	ANNs	0.47	0.96	-0.025	0.12	0.99	0.96	0.47	0.96	-0.036	0.12	0.99	0.96
	SVM	0.50	0.95	-0.002	0.13	0.99	0.95	0.50	0.96	-0.018	0.13	0.99	0.95
اسماعیل آباد Esmaeel Abad	RF	0.42	0.96	-0.016	0.12	0.99	0.96	0.44	0.96	-0.030	0.12	0.99	0.96
	ANNs	0.42	0.96	-0.012	0.12	0.99	0.96	0.41	0.97	-0.030	0.11	0.99	0.96
	SVM	0.47	0.96	0.047	0.13	0.99	0.95	0.47	0.96	0.026	0.13	0.99	0.95
بوئین زهرا Buinzahra	RF	0.44	0.96	-0.006	0.11	0.99	0.96	0.44	0.96	-0.005	0.12	0.99	0.96
	ANNs	0.43	0.96	-0.010	0.11	0.99	0.96	0.45	0.96	0.000	0.11	0.99	0.96
	SVM	0.48	0.96	0.018	0.12	0.99	0.95	0.47	0.96	0.024	0.12	0.99	0.95
تاکستان Takestan	RF	0.44	0.96	-0.003	0.12	0.99	0.96	0.48	0.96	0.001	0.13	0.99	0.96
	ANNs	0.43	0.96	0.000	0.11	0.99	0.96	0.43	0.97	0.003	0.11	0.99	0.96
	SVM	0.45	0.96	0.023	0.12	0.99	0.96	0.46	0.96	0.024	0.12	0.99	0.96
رازمیان Razmian	RF	0.45	0.95	-0.010	0.13	0.99	0.95	0.46	0.95	-0.018	0.14	0.99	0.95
	ANNs	0.47	0.95	-0.012	0.14	0.99	0.95	0.47	0.95	-0.045	0.14	0.99	0.95
	SVM	0.49	0.94	0.025	0.14	0.98	0.94	0.48	0.95	-0.010	0.14	0.98	0.94
سیردان Sirdan	RF	0.45	0.94	-0.001	0.15	0.98	0.94	0.47	0.93	-0.012	0.16	0.98	0.93
	ANNs	0.43	0.95	-0.016	0.14	0.99	0.95	0.43	0.95	-0.044	0.14	0.99	0.95
	SVM	0.46	0.94	0.033	0.15	0.98	0.94	0.47	0.94	-0.005	0.15	0.98	0.94
قزوین Qazvin	RF	0.42	0.96	-0.005	0.12	0.99	0.96	0.43	0.96	-0.012	0.12	0.99	0.96
	ANNs	0.40	0.97	0.001	0.11	0.99	0.97	0.40	0.97	-0.004	0.11	0.99	0.97
	SVM	0.44	0.96	0.017	0.12	0.99	0.96	0.44	0.96	0.005	0.12	0.99	0.96
کوهین Kuhin	RF	0.44	0.95	0.002	0.13	0.99	0.95	0.48	0.95	0.001	0.15	0.98	0.94
	ANNs	0.41	0.96	-0.005	0.13	0.99	0.96	0.41	0.96	0.001	0.13	0.99	0.96
	SVM	0.47	0.95	0.040	0.14	0.98	0.94	0.48	0.95	0.048	0.15	0.98	0.94
معلم کلایه Moalemkelayeh	RF	0.43	0.95	-0.005	0.14	0.99	0.95	0.46	0.95	-0.028	0.14	0.98	0.94
	ANNs	0.42	0.95	-0.008	0.13	0.99	0.95	0.42	0.95	-0.024	0.13	0.99	0.95
	SVM	0.45	0.95	0.032	0.14	0.98	0.94	0.44	0.95	0.008	0.14	0.99	0.95
ایستگاه‌های کلیماتولوژی Climatology stations													
		RMSE	R^2	MBE	NRMSE	d	EF	RMSE	R^2	MBE	NRMSE	d	EF
مگسال Megsal	RF	0.39	0.97	-0.002	0.12	0.99	0.97	0.42	0.96	0.012	0.13	0.99	0.96
	ANNs	0.41	0.96	-0.015	0.12	0.99	0.96	0.42	0.96	-0.014	0.13	0.99	0.96
	SVM	0.45	0.95	0.035	0.13	0.99	0.96	0.46	0.95	0.043	0.14	0.99	0.96
نیروگاه رجایی Rajai power plant	RF	0.33	0.97	0.001	0.09	0.99	0.97	0.35	0.97	0.001	0.10	0.99	0.97
	ANNs	0.35	0.97	0.020	0.10	0.99	0.97	0.35	0.97	0.022	0.10	0.99	0.97
	SVM	0.39	0.97	0.014	0.11	0.99	0.96	0.38	0.97	0.003	0.11	0.99	0.97
البرز Alborz	RF	0.42	0.97	-0.011	0.11	0.99	0.96	0.44	0.96	-0.022	0.12	0.99	0.96
	ANNs	0.39	0.97	-0.009	0.11	0.99	0.97	0.40	0.97	0.000	0.11	0.99	0.97
	SVM	0.46	0.96	-0.019	0.13	0.99	0.96	0.46	0.96	-0.004	0.13	0.99	0.96
باغ کوثر Bagh kosar	RF	0.38	0.97	-0.003	0.11	0.99	0.97	0.44	0.96	0.003	0.13	0.99	0.96
	ANNs	0.42	0.96	-0.014	0.13	0.99	0.96	0.44	0.96	0.000	0.13	0.99	0.96
	SVM	0.45	0.96	0.109	0.14	0.99	0.95	0.48	0.96	0.121	0.14	0.99	0.95
جیرکول Jirkol	RF	0.43	0.96	0.002	0.13	0.99	0.96	0.44	0.95	0.012	0.14	0.99	0.95
	ANNs	0.46	0.95	-0.007	0.14	0.99	0.95	0.46	0.95	0.014	0.14	0.99	0.95
	SVM	0.49	0.94	-0.017	0.15	0.99	0.94	0.49	0.94	0.004	0.15	0.99	0.94



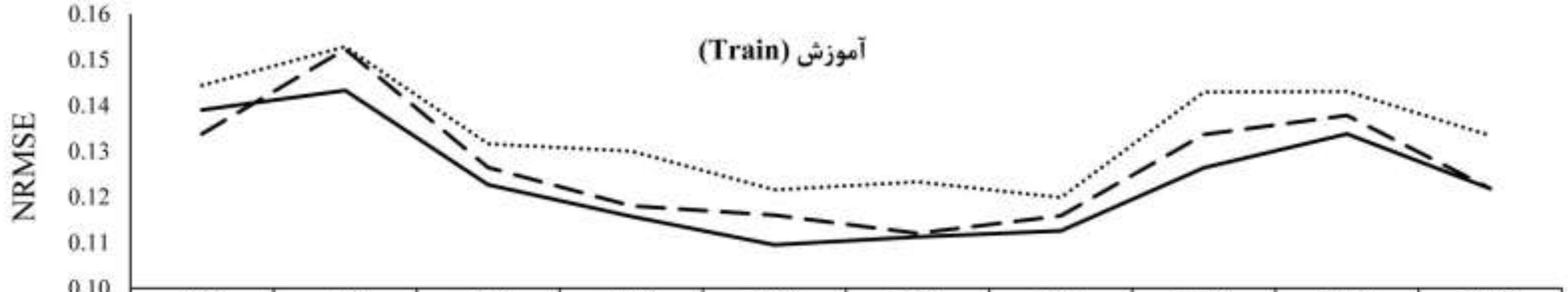
شکل ۲- مقایسه NRMSE سه روش در برآورد ET_0 در مرحله آزمون

Figure 2- Comparison the NRMSE of three methods for estimating of ET_0 in the test steps

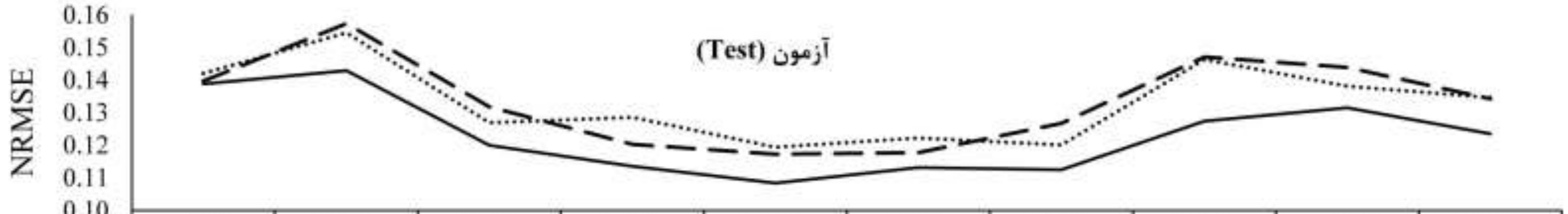
نتایج نشان می‌دهد (جدول ۵) که دقت سه روش به‌کاررفته در این پژوهش در ایستگاه نیروگاه رجایی نسبت به سایر ایستگاه‌های کلیماتولوژی، در هر دو مرحله آموزش و آزمون بیشتر می‌باشد، و NRMSE روش‌های بکار رفته در این ایستگاه نسبت به بقیه کمتر می‌باشد. و همچنین دقت تخمین در ایستگاه جیرکول نسبت به سایر ایستگاه‌های کلیماتولوژی در هر دو مرحله آموزش و آزمون کمتر و میانگین خطای بیشتری دارد. افزایش ارتفاع سبب کاهش دما و به‌تبع آن کاهش مقدار ET_0 و افزایش رطوبت نسبی می‌شود (Raziei *et al.*, 2005). در ایستگاه جیرکول به علت افزایش ارتفاع و کاهش مقدار ET_0،کاهش دقت تخمین در این ایستگاه منطقی می‌باشد. اما در ایستگاه نیروگاه رجایی که با افزایش ارتفاع، دقت تخمین افزایش‌یافته است نیاز به بررسی بیشتری می‌باشد. در این منطقه به علت فعالیت نیروگاه رجایی و تولید فراوان آلاینده‌های مضر که از ترسیب آن‌ها موجب آلودگی آب سطحی، خاک و گیاه می‌شود. حضور این آلاینده‌ها باعث باران اسیدی و درنتیجه نابودی پوشش‌های گیاهی و ساختار خاک می‌شود (Rai *et al.*, 2011). در این منطقه باغات و مزارع کشاورزی در جنوب، جنوب شرقی، جنوب غربی این نیروگاه قرار دارند، از طرفی جهت باد غالب در استان قزوین به سمت جنوب شرقی بوده (Shahryar *et al.*, 2019) و بیشترین میزان ترسیب آلاینده‌ها در این جهت می‌باشد که باعث افزایش ET_0 و کاهش رطوبت نسبی و همچنین خسارت به پوشش گیاهی می‌شود. این مسئله در جدول ۱ مشاهده می‌شود، میانگین رطوبت نسبی در ایستگاه نیروگاه رجایی با وجود ارتفاع بالاتر، کمتر از سایر ایستگاه‌ها است. کیخسروی و همکاران (Keikhosravi *et al.*, 2014) در ایستگاه‌های سینوپتیک استان کرمانشاه ET_0 را محاسبه کردند، و نشان دادند که عامل ارتفاع باعث اختلاف در عناصر اقلیمی و به‌تبع آن مقادیر ET_0 شده است.

در پژوهش حاضر با افزایش ارتفاع از ۱۳۰۰ متر به بالا، میانگین خطا به‌طور قابل‌توجهی افزایش می‌یابد (شکل ۳). علت آن‌هم احتمالاً به دلیل تغییرات دمای هوا با افزایش رطوبت و تأثیر آن بر مقدار ET_0 می‌باشد. ساعت آفتابی که تابعی از عرض جغرافیایی می‌باشد، در فصول مختلف سال متفاوت است. به‌طوری‌که در تمامی فصول با افزایش عرض جغرافیایی از میزان تشعشع ورودی به زمین کاسته می‌شود. بنابراین انتظار می‌رود که مقدار ET_0 در عرض‌های جغرافیایی بالاتر، کمتر از عرض‌های جغرافیایی پایین‌تر به دست آید، ازجمله دلایل این امر می‌تواند مرتبط با در نظر گرفتن عواملی مانند عرض جغرافیایی و ارتفاع از سطح دریا در محاسبات مقدار ET_0 باشد در اثر تغییرات اقلیمی، مقدار ET_0 نه‌تنها بر اساس دمای هوا بلکه بر اساس سایر متغیرهای آب و هوایی دست‌خوش تغییرات قرار می‌گیرد (Yang *et al.*, 2017). بنابراین نتایج تحقیق حاضر نشان می‌دهند که اثر ارتفاع بر تغییرات مقدار ET_0 تحت شرایط آب و هوایی متغیر، نامشخص است. کنگ و همکاران (Kang *et al.*, 2021) نتایج مشابه‌ایی ارائه کردند.

با توجه به جدول ۵، مقادیر آماره MBE در مرحله آموزش و آزمون اکثراً مقادیر کوچک و نزدیک به صفر می‌باشد. بنابراین با توجه به کم بودن مقادیر آماره MBE، نمی‌توان گفت مدل کم‌برآوردی یا بیش برآوردی قابل‌توجهی داشته است و درنتیجه می‌توان از آن چشم‌پوشی کرد (Singh et al., 2020). در جدول۵ مشاهده می‌شود آماره EF در همه ایستگاه‌ها در هر دو مرحله آموزش و آزمون بالای ۰/۹۴ می‌باشد و همچنین شاخص d هم در همه ایستگاه‌ها در هر دو مرحله آموزش و آزمون بین ۰/۹۸-۰/۹۹ می‌باشد،که نشان از صحت و کارایی بالای روش‌های استفاده در این پژوهش در تخمین مقدار ET_0 می‌باشد. از طرفی فاکتورهای مختلفی بر عملکرد روش‌های یادگیری ماشین مانند الگوریتم آموزش ANNs، یا تنظیمات فرا پارامترهای روش‌های مورداستفاده، دخیل هستند (Ferreira *et al.*, 2019). بنابراین در این تحقیق انتخاب عوامل مؤثر در تخمین مقدار ET_0 مناسب بوده است که مقادیر آماره‌های EF و d بالا بوده است.



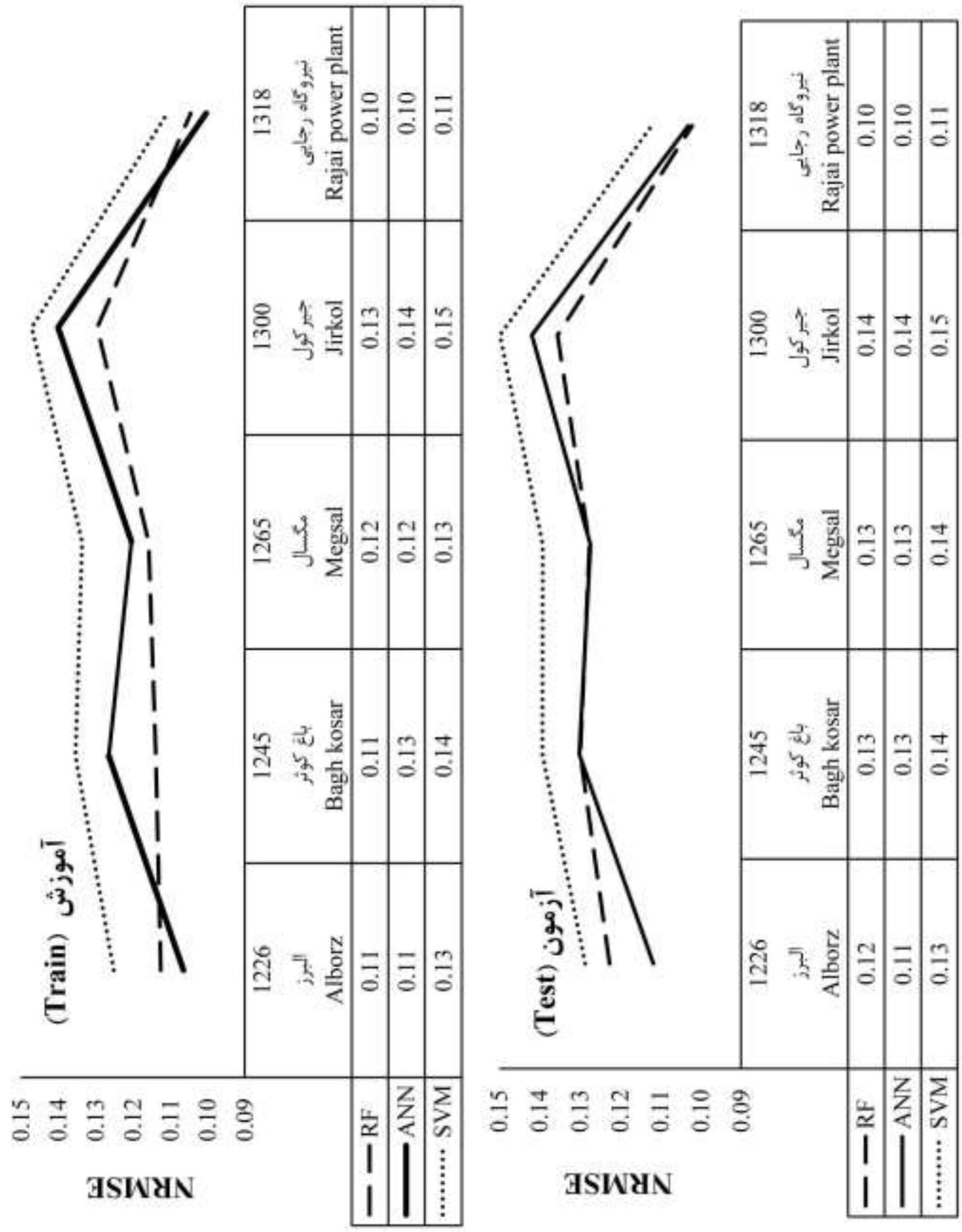
	982 رازمیان Razmian	1100 سیردان Sirdan	1226 آبیک Abyek	1249 اسماعیل آباد Esmaeel Abad	1280 قزوین Qazvin	1282 بوئین زهرا Buinzahra	1284 تاکستان Takestan	1549 کوهین Kuhin	1630 معلم کلایه Moalemkelayeh	2035 آوج Avaj
RF	0.13	0.15	0.13	0.12	0.12	0.11	0.12	0.13	0.14	0.12
ANN	0.14	0.14	0.12	0.12	0.11	0.11	0.11	0.13	0.13	0.12
SVM	0.14	0.15	0.13	0.13	0.12	0.12	0.12	0.14	0.14	0.13



	982 رازمیان Razmian	1100 سیردان Sirdan	1226 آبیک Abyek	1249 اسماعیل آباد Esmaeel Abad	1280 قزوین Qazvin	1282 بوئین زهرا Buinzahra	1284 تاکستان Takestan	1549 کوهین Kuhin	1630 معلم کلایه Moalemkelayeh	2035 آوج Avaj
RF	0.14	0.16	0.13	0.12	0.12	0.12	0.13	0.15	0.14	0.13
ANN	0.14	0.14	0.12	0.11	0.11	0.11	0.11	0.13	0.13	0.12
SVM	0.14	0.15	0.13	0.13	0.12	0.12	0.12	0.15	0.14	0.13

شکل ۳- تغییرات خطا با افزایش ارتفاع در تخمین ET_0 در ایستگاه‌های سینوپتیک

Figure 3- Error changes with increasing altitude in ET_0 estimation



	1226 البرز Alborz	1245 باغ کوثر Bagh kosar	1265 مگسال Megsal	1300 جیرکول Jirkol	1318 نیروگاه رجایی Rajai power plant
RF	0.11	0.11	0.12	0.13	0.10
ANN	0.11	0.13	0.12	0.14	0.10
SVM	0.13	0.14	0.13	0.15	0.11

	1226 البرز Alborz	1245 باغ کوثر Bagh kosar	1265 مگسال Megsal	1300 جیرکول Jirkol	1318 نیروگاه رجایی Rajai power plant
RF	0.12	0.13	0.13	0.14	0.10
ANN	0.11	0.13	0.13	0.14	0.10
SVM	0.13	0.14	0.14	0.15	0.11

شکل ٤- تغییرات خطا با افزایش ارتفاع در تخمین ET0 در ایستگاه‌های کلیماتولوژی
Figure 4- Error changes with increasing altitude in ET0 estimation

نتیجه‌گیری

نتایج پژوهش حاضر نشان داد مقادیر ET_0 سامانه نیاز آب، همبستگی خوبی با مقادیر ET_0 تخمینی روش‌های ANNs، RF، و SVM دارد. در این پژوهش دقت تخمین مقدار ET_0 با دو روش ANNs و RF نزدیک به هم و نسبت به روش SVM بالاتر بوده است. نتایج نشان داد، مقدار ET_0 نه‌تنها بر اساس دمای هوا بلکه ممکن است تحت تأثیر عامل‌های دیگر مانند آلودگی هوا تغییر کند، همچنین به‌شدت تحت تأثیر شرایط منطقه‌ای مانند توپوگرافی و ارتفاع می‌باشد. نتایج این تحقیق علاوه بر بررسی بهتر ET_0، به شناخت عوامل تأثیرگذارتر در هر منطقه کمک فراوان می‌کند و می‌توان در مناطق با اقلیم‌های مشابه از آن استفاده کرد. به‌عنوان‌مثال در منطقه مطالعاتی حاضر، مشخص شد که میانگین دمای هوا بیشتر از سایر پارمترهای اقلیمی بر ET_0 تأثیر می‌گذارد. با درک این مسئله، چنانچه دمای هوای یک منطقه زیاد شود، باعث افزایش شدید در

مقدار ET_0 و متعاقباً افزایش نیاز آبی گیاهان خواهد شد. توجه به این نکته خواهد توانست از بروز تنش آبی و کاهش احتمالی محصول جلوگیری کند. با توجه به متفاوت بودن ارتفاع و عرض جغرافیایی ایستگاه‌ها، و از طرفی شرایط اقلیمی متفاوت در این ایستگاه‌ها، می‌توان چنین نتیجه گرفت که نقش اقلیم در تخمین ET_0 مهم‌تر از نقش ارتفاع و عرض جغرافیایی است. نتیجه نهایی اینکه الگوریتم‌های مختلف روش‌های داده‌کاوی، در شرایطی منطقه‌ایی مانند سیردان و رازمیان که داده‌های اقلیمی محدودی در دسترس می‌باشد، می‌توانند در مطالعات آینده برای توسعه مدل‌های تخمینی ET_0 مناسب باشند.

منابع

1. Adab, H., Morbidelli, R., Saltalippi, C. Moradian, M., & Ghalhari, G.A.F, (2020). Machine learning to estimate surface soil moisture from remote sensing data. *Water 12*(11): 3223. https://doi.org/10.3390/w12113223.
2. Breiman, L. (2001). Random forests. *Machine Learn*. 45: 5–32. https://doi.org/10.1023/A:1010933404324.
3. Chen, Z., Zhu, Z., Jiang, H., & Sun, S. (2020). Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. *Journal of Hydrology* 591: 125286. https://doi.org/10.1016/j.jhydrol.2020.125286.
4. Dos Santos Farias, D.B., Althoff, D., Rodrigues, L.N., & Filgueiras, R. (2020). Performance evaluation of numerical and machine learning methods in estimating reference evapotranspiration in a Brazilian agricultural frontier. *Theoretical and Applied Climatology* 142: 1481-1492. https://doi.org/10.1007/s00704-020-03380-4.
5. Fazeli Khiavi, A., Salahi, B., & Goodarzi, M. (2020). Assessment effects of climate change on changes in potential evapotranspiration in the Moghan Plain by rcps. *Watershed Engineering and Management* 12: 977-993. (In Persian). https://doi.org/10.22092/ijwmse.2019.126245.1649.
6. Feng, K., & Tian, J. (2021). Forecasting reference evapotranspiration using data mining and limited climatic data. *European Journal of Remote Sensing* 54: 363-371. https://doi.org/10.1080/22797254.2020.1801355.
7. Ferreira, L.B., da Cunha, F.F., de Oliveira, R.A., & Fernandes Filho, E.I. (2019). Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM–A new approach. *Journal of Hydrology* 572: 556-570. https://doi.org/10.1016/j.jhydrol.2019.03.028.
8. Gavili, S., Sanikhani, H., Kisi, O., & Mahmoudi, M.H. (2018). Evaluation of several soft computing methods in monthly evapotranspiration modelling. *Meteorological Applications* 25: 128-138. https://doi.org/10.1002/met.1676.
9. Gopinathan, K.K. (1988). A general formula for computing the coefficients of the correlation connecting global solar radiation to sunshine duration. *Solar Energy* 41: 499-502. https://doi.org/10.1016/0038-092X(88)90052-7.
10. Hagan, M.T., Demuth, H.B., & Beale, M.H. (1996). *Neural Network.* Design PWS Publishing Co.
11. Hocking, R.R. (2013). *Methods and applications of linear models: regression and the analysis of variance*, John Wiley & Sons.
12. Jing, W., Yaseen, Z.M., Shahid, S., Saggi, M.K., Tao, H., Kisi, O., Salih, S.Q., Al-Ansari, N., & Chau, K.-W. (2019). Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions. *Engineering Applications of Computational Fluid Mechanics* 13: 811-823. https://doi.org/10.1080/19942060.2019.1645045.
13. Kang, T., Li, Z., & Gao, Y. (2021). Spatiotemporal variations of reference evapotranspiration and its determining Climatic factors in the Taihang Mountains, China. *Water* 13: 3145. https://doi.org/10.3390/w13213145.
14. Karimipour, A., & Banitalebi, G. (2020). Sensitivity analysis of meteorological data in estimating reference evapotranspiration with the minimum data using wavelet-neuro-fuzzy, ANN and ANFIS models. *Journal of Soil and Water Resources Conservation 9*(3):47-72. (In Persian).
15. Keikhosravi, G., Rezaee, A., Mohamadi, Z., & Baghaee, M. (2014). *The Estimation of Reference Evapotranspiration in (reference grass) 5 synoptic Station province of Kermanshah with using REF-ET Model. National conference of new ideas in sustainable agriculture.* Islamic Azad University, Borujerd branch 1-18, (In Persian).
16. Majozi, N.P., Mannaerts, C.M., Ramoelo, A., Mathieu, R., & Verhoef, W. (2021). Uncertainty and sensitivity analysis of a remote-sensing-based penman–Monteith model to meteorological and land surface input variables. *Remote Sensing* 13: 882. https://doi.org/10.3390/rs13050882.
17. Mattar, M.A., (2018). Using gene expression programming in monthly reference evapotranspiration modeling: a case study in Egypt. *Agricultural Water Management* 198: 28-38. https://doi.org/10.1016/j.agwat.2017.12.017.
18. Monteith, J. (1965). *The state andmovement of water in living organisms*. In: 19th Symposia of the Society for Experimental Biology. Cambridge University Press, London 205–234.
19. Ndiaye, P.M., Bodian, A., Diop, L., Deme, A., Dezetter, A., Djaman, K., & Ogilvie, A. (2020). Trend and sensitivity analysis of reference evapotranspiration in the Senegal river basin using NASA meteorological data. *Water* 12: 1957. https://doi.org/10.3390/w12071957.

20. Panaitescu, L., Ilie, C., Lungu, M., Popescu, M., Lungu, D., & Nita, S. (2014). Modern approach to the phenomenon of drought and aridity in Central and South Dobrudja. *Journal of Environmental Protection and Ecology* 15: 110-122.
21. Picton, P. (2000). *Neural Networks*, 2nd edn. Palgrave, New York.
22. Rai, R., Rajput, M., Agrawal, M., & Agrawal, S. (2011). Gaseous air pollutants: a review on current and future trends of emissions and impact on agriculture. *Journal of Scientific Research 55*(771): 1.
23. Raza, A., Shoaib, M., Khan, A., Baig, F., Faiz, M.A., & Khan, M.M. (2020). Application of non-conventional soft computing approaches for estimation of reference evapotranspiration in various climatic regions. *Theoretical and Applied Climatology* 139: 1459-1477. https://doi.org/10.1007/s00704-019-03007-3.
24. Raziei, T., Daneshkar Arasteh, P., & Saghafian, B. (2005). Annual rainfall trend analysis in arid and semi-arid regions of central and eastern Iran. *Water and Wastewater* 54: 73-81 (In Persian).
25. Saggi, M.K., & Jain, S. (2019). Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. *Computers and Electronics in Agriculture* 156: 387-398. https://doi.org/10.1016/j.compag.2018.11.031.
26. Shahryar, F., Gandomkar, A., & Hashempour, R. (2019). Optimal locating of the new towns in Qazvin Province based on climatic parameters. *Geography and Environmental Planning* 29: 19-34. (In Persian). https://doi.org/10.22108/gep.2018.98275.0.
27. Shiri, J., Marti, P., Karimi, S., & Landeras, G. (2019). Data splitting strategies for improving data driven models for reference evapotranspiration estimation among similar stations. *Computers and Electronics in Agriculture* 162: 70-81. https://doi.org/10.1016/j.compag.2019.03.030.
28. Singh, A., Haghverdi, A., Öztürk, H.S., & Durner, W. (2020). Developing pseudo continuous pedotransfer functions for international soils measured with the evaporation method and the HYPROP system: I. The soil water retention curve. *Water* 12: 3425. https://doi.org/10.3390/w12123425.
29. Sandhu, R., & Irmak, S. (2020). Performance assessment of hybrid-maize model for rainfed, limited and full irrigation conditions. *Agricultural Water Management* 242: 106402. https://doi.org/10.1016/j.agwat.2020.106402.
30. Tabari, H., Martinez, C., Ezani, A., & Hosseinzadeh Talaee, P. (2013). Applicability of support vector machines and adaptive neurofuzzy inference system for modeling potato crop evapotranspiration. *Irrigation science* 31: 575-588. https://doi.org/10.1007/s00271-012-0332-6.
31. Üneş, F., Kaya, Y.Z., & Mamak, M. (2020). Daily reference evapotranspiration prediction based on climatic conditions applying different data mining techniques and empirical equations. *Theoretical and Applied Climatology* 141: 763-773. https://doi.org/10.1007/s00704-020-03225-0.
32. Vapnik, V.N. (2000). *The nature of statistical learning theory, ser. Statistics for engineering and information science*, Springer, New York, 21:1003–1008.
33. Willmott, C.J., Robeson, S.M., & Matsuura, K. (2012). A refined index of model performance. International Journal of climatology 32: 2088-2094. https://doi.org/10.1002/joc.2419.
34. Yang, L., Feng, Q., Li, C., Si, J., Wen, X., & Yin, Z. (2017). Detecting climate variability impacts on reference and actual evapotranspiration in the Taohe River Basin, NW China. *Hydrology Research* 48: 596-612. https://doi.org/10.2166/nh.2016.252.